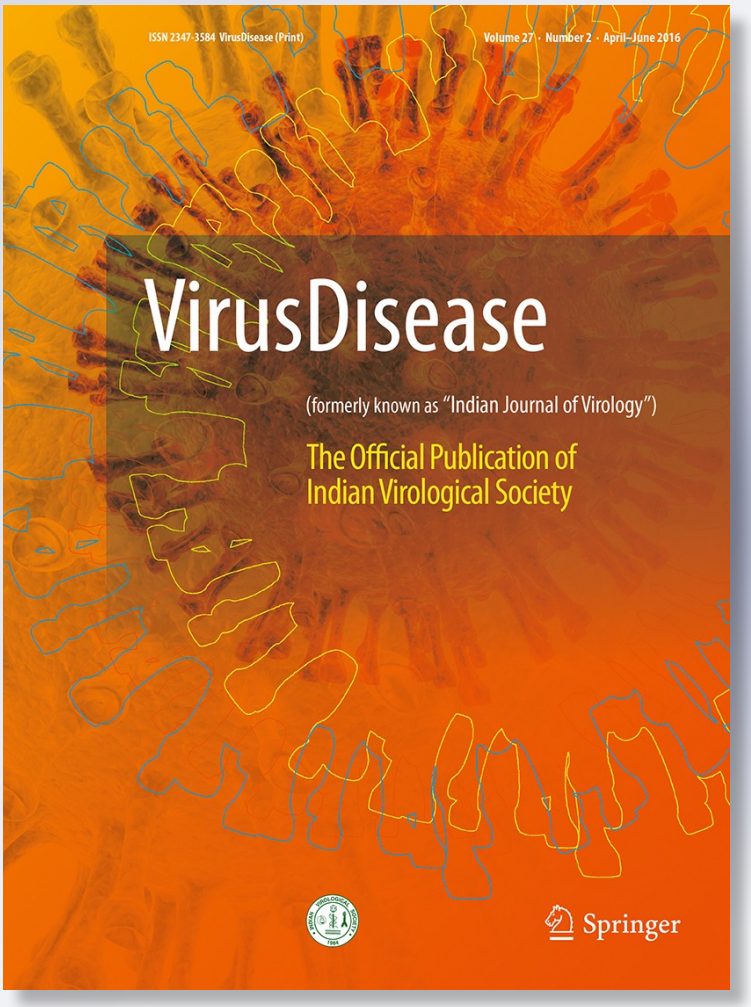
Ebolavirus evolves in human to minimize the detection by immune cells by accumulating adaptive mutations

Arunachalam Ramaiah & Vaithilingaraja Arumugaswami

VirusDisease

ISSN 2347-3584
Volume 27
Number 2

VirusDis. (2016) 27:136-144
DOI 10.1007/s13337-016-0305-0



Springer

VirusDis. (April–June 2016) 27(2):136–144
DOI 10.1007/s13337-016-0305-0

ORIGINAL ARTICLE

Ebolavirus evolves in human to minimize the detection by immune cells by accumulating adaptive mutations

Arunachalam Ramaiah[1] · Vaithilingaraja Arumugaswami[2,3,4]

Received: 25 January 2016 / Accepted: 1 February 2016 / Published online: 22 February 2016

Abstract The current outbreak of *Zaire ebolavirus* (EBOV) lasted longer than the previous outbreaks and there is as yet no proven treatment or vaccine available. Understanding host immune pressure and associated EBOV immune evasion that drive the evolution of EBOV is vital for diagnosis as well as designing a highly effective vaccine. The aim of this study was to deduce adaptive selection pressure acting on each amino acid sites of EBOV responsible for the recent 2014 outbreak. Multiple statistical methods employed in the study include SLAC, FEL, REL, IFEL, FUBAR and MEME. Results show that a total of 11 amino acid sites from sGP and ssGP, and 14 sites from NP, VP40, VP24 and L proteins were inferred as positively and negatively selected, respectively. Overall, the function of 11 out of 25 amino acid sites under selection pressure exactly found to be involved in T cell and B-cell epitopes. We identified that the EBOV had evolved through purifying selection pressure, which is a predictor that is known to aid the virus to adapt better to the human host and subsequently reduce the efficiency of existing immunity. Furthermore, computational RNA structure prediction showed that the three synonymous nucleotide mutations in NP gene altered the RNA secondary structure and optimal base-pairing energy, implicating a possible effect on genome replication. Here, we have provided evidence that the EBOV strains involved in the recent 2014 outbreak have evolved to minimize the detection by T and B cells by accumulating adaptive mutations to increase the survival fitness.

Keywords *Zaire ebolavirus* · Natural selection · Immune evasion · Vaccine

Introduction

Lethal hemorrhagic fever is caused by a single-stranded negative-sense RNA virus known as *Zaire ebolavirus* (EBOV), which is one among five species of genus *Ebolavirus*. The genus *Ebolavirus* belongs to the family of *Filoviridae*. The recent 2014 EBOV outbreak in West Africa led to deaths up to 90 % of infected humans and has now spread to Spain and the United States of America [28]. Typically, EBOV infected patients die within 1–2 weeks and is dependent on the viral strain and host species. Due to the acute effect on the host immune system, recovery time of EBOV infected individual is a lengthy and complex process. EBOV targets the dendritic cells, monocytes and macrophages during the early stages of infection, while, non-lymphocytic cell types at the later stages [18, 19]. As there are no effective vaccines and anti-viral therapies available for EBOV, serious efforts are being taken to develop the same in order to prevent infection and cure the disease [7]. Typically, viruses have evolved to prevent immune responses in a selective and regulated way, which

Electronic supplementary material The online version of this article (doi:10.1007/s13337-016-0305-0) contains supplementary material, which is available to authorized users.

✉ Arunachalam Ramaiah
arunachalamphd@gmail.com

[1] Centre for Infectious Disease Research, Indian Institute of Science, Bangalore 560 012, India

[2] Board of Governors Regenerative Medicine Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

[3] Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

[4] Department of Surgery, University of California at Los Angeles, Los Angeles, CA 90095, USA

facilitates dominated virus replication and progeny transmission [18]. Further advancement in the understanding of host-viral interactions is pertinent and beneficial to subsequently limiting and eliminating EBOV infections. Also, it is interesting to test the hypothesis that pathogens use antigenic variation as one among various mechanisms to escape from human immune responses [5]. Thus, the goals of this study was to deduce the adaptive selection pressure acting on each amino acid sites of EBOV responsible for the recent 2014 outbreak, as it is very crucial in the effective design of vaccines, and to infer the function of amino acid sites under selection pressures.

Materials and methods

Sequence selection

The EBOV genome consists of seven genes such as nucleoprotein (NP), the membrane-associated viral protein (VP24), polymerase matrix viral proteins (VP30 and VP35), matrix viral protein (VP40), the RNA-dependent RNA polymerase (L), and the glycoprotein (GP), which are transcribed into mRNA, resulting in the production of at least ten proteins [15, 18]. Notably, gene GP encodes two proteins, namely secreted glycoprotein (sGP) and small secreted glycoprotein (ssGP), as a result of mRNA editing. Total of 97, 98, 100, 98, 98, 98, 97, 100 and 100 protein-coding sequences of the EBOV responsible for 2014 outbreak were retrieved from the GenBank (dated 06-Nov-2014) for the NP, VP24, VP30, VP35, VP40, L, GP, sGP and ssGP, respectively. It should be noted that, gene GP has not been included as it has several non-sense stop codons at the following amino acid positions 365, 430, 479, 481, 489, 492, 503, 547, 554, 558, 573, 619 and 662. The nucleotide sequence 1–895 of sGP overlaps with the entire ssGP in a different reading frame. The acquired sequences were selected accordingly using the following criteria: (1) strains responsible for the 2014 outbreak and (2) sequences derived from the same strain, artificial sequences and incomplete sequences were removed. The finally considered 2014 Ebola strains and their GenBank accession numbers listed in the Supplementary Table 1.

Natural selection analysis

The selection pressures operating on each codon/amino acid site of EBOV were detected by computing 'ω', the ratio between non-synonymous (dN) and synonymous (dS) substitutions. The selection analyses were made using the following statistical methods from datamonkey [8]: four independent maximum likelihood methods that include single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), random effects likelihood (REL) [25], internal fixed effects likelihood (IFEL) [24], single bayesian method namely fast unconstrained bayesian approximation (FUBAR) [22], single branch-site method namely mixed effects model of episodic selection (MEME) [26]. These methods are described in detail [1, 2, 22, 24–26]. In brief, the coding regions of EBOV were aligned using MEGA6 [30] and then confirmed sequences were used to input in datamonkey. The model selection test was employed to find out the best model for each gene/protein (Table 1). The 'ω' ratio of amino acid sites under selection pressures either with p values ≤ 0.05 (SLAC, FEL, IFEL and MEME) or posterior probability ≥ 0.9 (FUBAR) or Bayes factors ≥ 50 (REL) were considered as statistically significant.

RNA structure and amino acids function prediction

Using the mfold web server [33] in default settings, RNA secondary structure was predicted for the segments, which had mutations that could be identified by multiple statistical methods. The function of the amino acid sites under selection pressure was inferred using functionally known immune epitopes from the Immune Epitope Database (IEDB) [31].

Results

Amino acids under positive Darwinian selection

The results on positive (diversifying) and negative (purifying) selection pressures acting on each amino acid site of all the proteins encoded by EBOV genomes were shown in Tables 1, 2, and 4. Overall, we were able to identify 11 amino acid sites in the sGP and ssGP proteins of EBOV that underwent positive Darwinian selection. All of the 11 positively selected sites were identified by a single method with statistical significance (Tables 2, 4); therefore, there was no identical positively selected site shown by other methods used in this study. Notably, no site was inferred as positively selected for NP, VP35, VP40, VP24 and L proteins (Table 2). VP30 has not been considered for the analysis as minimum three non-identical sequences are required, but only two non-identical sequences exist. Sites 130, 134, and 136 of ssGP overlap with sGP at identical positions; however, these were considered as different sites since they came from individual (paralogues) proteins.

Amino acids under negative selection

The amino acid sites of EBOV under purifying selection pressures are relatively higher (Tables 1, 2). The present

Table 1 Summary of diversifying and purifying selection pressure acting on each amino acid of EBOV

Protein	Total no. of amino acid sites	Model	AIC value	SLAC	FEL	REL	IFEL	FUBAR	MEME	$\omega > 1/\omega < 1$ (total no. of sites)[a]	No. of positively/negatively selected sites with $p \leq 0.05$ (or) BF ≥ 50 (or) Posterior Prob ≥ 0.9[b]	No. of statistically reliable positively/negatively selected sites identified by more than one method[c]
NP	739	F81	6366.79	–/1	–/4	–/–	–/–	–/4	–/–	–/4	–/4	–/3
VP35	340	F81	2859.35	–/–	–/1	–/–	–/–	–/–	–/–	–/1	–/–	–/–
VP40	326	F81	2781.83	–/–	–/–	–/–	–/–	–/1	–/–	–/1	–/1	–/–
sGP	364	F81	3064.65	–/–	–/–	8/–	–/–	–/–	–/–	8/–	8/–	–/–
ssGP	297	F81	2511.90	–/–	–/–	3/–	–/–	–/–	–/–	3/–	3/–	–/–
VP30	288	–	–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–
VP24	251	F81	2136.78	–/–	–/1	–/–	–/–	–/1	–/–	–/2	–/2	–/–
L	2212	HKY85	18483.96	–/–	–/3	–/–	–/–	–/7	–/–	–/7	–/7	–/–

[a] The total number of positively and negatively selected sites were observed from different methods for each gene/protein

[b] Among the total number of positively and negatively selected sites, amino acid sites with statistical significant levels [$p \leq 0.05$ (SLAC, FEL, IFEL and MEME) or Bayes factors $\geq$50 (REL) or posterior probability $\geq$0.9 (FUBAR)] were indicated

[c] The total number of statistically reliable positively and negatively selected sites for each gene observed by more than one method was indicated

Table 2 Details of statistically reliable diversifying and purifying selected amino acid sites of EBOV

Protein	Method	Total no. of statistically reliable positively selected sites		Position of diversifying selected amino acids	Total no. of statistically reliable negatively selected sites		Position of purifying selected amino acids
NP	FEL	0	0	0	3	4	185, **341, 460, 624**
	FUBAR	0			4		
VP40	FUBAR	0	0	0	1	1	9
sGP	REL	8	8	11, 54, 130, 134, 136, 323, 324, 363	0	0	0
ssGP	REL	3	3	130, 134, 136	0	0	0
VP24	FUBAR	0	0	0	1	2	55, 133
	FEL	0			1		
L	FUBAR	0	0	0	7	7	121, 781, 813, 1461, 1625, 1854, 2135

There was no statistically reliable positively selected site found for NP, VP35, VP40, VP24 and L. On the other hand, no statistically reliable negatively selected site was found for sGP and ssGP. The positively and negatively selected common amino acid sites identified by more than one method in the datamonkey were denoted by bold. For instance, site **341** (bold) of NP was notified as a negatively selected common site that could be identified by both FEL and FUBAR

data shows that a total of 14 out of 15 negatively selected amino acid sites were inferred with statistical significance which were relatively higher in L (seven sites) protein than in NP, VP24 and VP40 proteins. However, none of the sites were inferred from sGP and ssGP proteins. Although a single negatively selected site (position 170) found in VP35, but was not statistically significant (FEL, *p value* 0.09). Moreover, sites 185 and 460 of NP have also been identified by FEL and SLAC methods, respectively, without statistical significant. As like NP, in the L protein also the amino acid positions 781, 1625 and 2135 have been identified by FEL, but with no reliable statistical *p-values* 0.07, 0.09 and 0.09, respectively.

Synonymous mutations alter the RNA secondary structure

We computationally predicted the secondary RNA structure of 3 of 11 NP protein-coding genes (obtained from accession #KJ660348; KM233044; KM034558) considered in the present study, which carry at least one of the three more reliable synonymous mutations. The comparison

results showed that the synonymous substitutions in the 3 codons of NPs altered the RNA secondary structures (Fig. 1a–c; Table 3) by generating mispaired stems and stem-loop interactions. Also, the energy dot blot analysis of these three structures indicated that the synonymous substitutions changed the base pairing and optimal energies (−670.4 to −665.9 kcal/mol) (Fig. 1d–f; Table 3).

Putative biological role of amino acids under selection

The biological function of amino acid sites under selection pressure was identified with known immune epitopes (Table 4). We observed that a total of 7 in sGP and 3 in ssGP of 11 positively selected amino acid sites were found

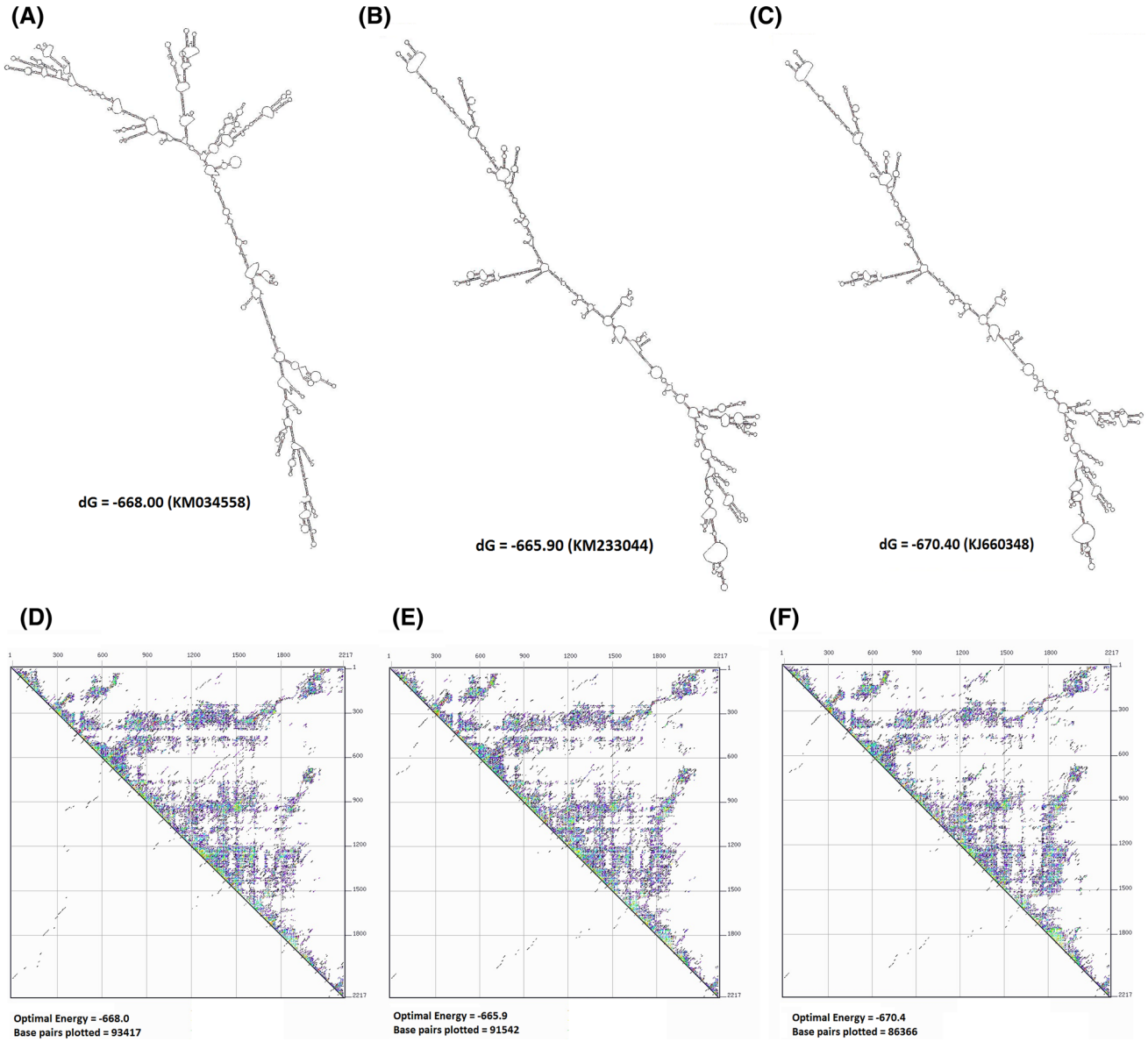


Fig. 1 Impact of synonymous mutations in stem-loop structures in protein-coding region of 3 NPs. **a–c** Predicted stem-loop structures of 3 NPs sequences using RNA mfold analysis. Importantly, the altered Watson–Crick base pair interactions in the 3 NPs structures leads to the variation in optimal energies, resulting in a altered RNA structures. The folding Gibbs free energy in kcal/mol for the predicted stem-loops was shown under each structure. **d–f** Energy dot plots for optimal and suboptimal folding of 3 NPs RNA structures. The upper triangle displays possible base pair combinations at various energy levels. The nucleotide positions of each base are displayed on the top axis and right axis of the upper triangle, whereas, the lower triangle shows the paired bases with optimal folding energy at 37 °C to form a stem-loop structure

Table 3 Summary of RNA structure predictions for 3 NP genes using mfold web server

More reliable amino acid position in NP protein	Codon	Synonymous nucleotide changes in 11 NP genes and the accession number of representative gene carrying the mutation		Accession number of NP gene used for RNA structure prediction	Optimal energy of predicted RNA structure
		Type-1 change	Type-2 change		
341	1021–1023	CAA (10) KM034558	CAG (1) KM233044	KM034558	−668.0
460	1378–1380	GAC (9) KM034558	GAT (2) KJ660348	KJ660348	−670.4
624	1870–1872	GAA (10) KM233044	GAG (1) KM034558	KM233044	−665.9

to be involved in T-cell epitopes (TCEs). In addition, 3 sGP sites (54, 323 and 324) are part of B-cell epitopes (BCEs) [3]. Notably, the adjacent site 55 was found to be involved in TCE, suggesting that site 54 might also be involved in immune recognition function. The epitope function of the sGP position 363 is not verified. Non-synonymous mutations in T-cell and B-cell epitopic regions of EBOV may be caused by the immune response mounted by the human host. These amino acid sites are also predicted to support EBOV to better adapt to the human and can reduce the efficiency of existing immunity.

The function for 4 out of 14 negatively selected amino acid sites of NP, VP40, VP24, and L were identified (Table 4). Amino acid sites 185 and 341 of NP protein found to be involved in TCEs, whereas, sites 460 and 624 are part of BCEs. However, possible involvement as epitope for remaining 10 sites of VP40, VP24 and L were unknown.

Discussion

The presence of total of 11 statistically reliable positively selected sites includes 8 and 3 sites in sGP and ssGP, respectively, indicate that these are an effect of molecular adaptation of the EBOV to human, which confers an evolutionary advantage to the virus. The GP looks to evolve faster when comparing to the other genes of EBOV [15]. It confirms the hypothesis that in the early stages of viral outbreaks, the virus adapts into a new host (e.g. human) will be dealing with a higher selection pressures [1, 2, 29]. Especially, codons/amino acids under positive Darwinian selection are subjected to stringent pressure [1, 10]. Notably, sGP protein is secreted by infected cells and is not present in virions. Previous study showed that to promote immune evasion, sGP serves as an antibody decoy or presents alternative non-neutralizing antibody epitopes for the humoral immune response [17, 20, 32]. Another study revealed that the 71 and 81 % of humoral immune response of symptomatic and asymptomatic individuals, respectively from Gabonese EBOV outbreaks mainly targets GP peptides [3]. Here, the reactive peptides located in linear immunogenic regions at amino acid positions 41–67 and 301–359 of GP are including the three positive selected sites 54, 323 and 324, identified in the present study. In context of the previous findings, these 3 positively selected sites of sGP expected to modulate or misdirect humoral immunity and aid the EBOV to adapt better in human host. Moreover, it is worth to note that the presented 3 (positions 54, 323 and 324) of 8 positively selected sites of sGP are involving in B-cell epitopes (BCEs) [3] and can be used as targets for developing new therapeutics.

The VP24 protein of EBOV found to be involved in blocking the production of interferon (IFN) in host resulting in attenuation of both type I (IFN alpha and beta) and type II (IFN gamma) signaling responses. In addition, VP24 prevents the nuclear accumulation of dimerized phosphorylated STAT-1 that participates in activation of both type I and type II signalling cascades [13, 23, 27]. It is possible that the negatively selected VP24 sites, 55 and 133, might aid the EBOV to become more pathogenic to human by inhibiting IFN response and blocking antigen presentation to T-cells [18]. Importantly, considering these two negatively selected VP24 amino acids include in the T-cell epitopes which can be used for vaccine design and testing.

The abundantly expressed matrix protein VP40 plays an important role in virion assembly and budding [12]. In the absence of other viral proteins, VP40 is also able to produce virus-like particles (VLPs) [3, 11, 14]. The NP and VP24 are crucial for viral replication and uncoating steps. L protein is known to be a part of the replicase complex. It associates with the ribonucleoprotein complex via VP35, which subsequently forms larger complex with VP40. Previous report also shows that strong humoral responses directed against NP and VP40 proteins [21]. The four amino acid sites of NP and VP40 proteins under negative selection pressure can be considered as main antibody targets for developing vaccine. In order to block the formation of the EBOV replication complex, seven negatively selected amino acid sites of L protein can be considered as promising target to develop antiviral drugs.

Table 4 Function of amino acid sites of EBOV under selection pressure

Protein	Statistically reliable amino acid sites	Method(s)	dN-dS/ω[a]	*p*-value/Posterior Prob./Bayes factor[b]	Functionally known Epitope length (from-to)	IEDB-Epitope ID[c]	Linear sequence(s)[d]	Function(s)
Positively selected amino acid sites								
sGP	11	REL	4.47	50.80	11–19	53414	**R**DRFKRTSF	TCE (HLA-A*24:02; HLA-B*07:02, 08:01, 15:01)
					7–15	91685	LQLP**R**DRFK	TCE (HLA-A*03:01, 11:01)
					4–12	92037	TGILQLP**R**D	TCE (HLA-A*02:01, 24:02)
	54#	REL	4.47	50.80	55–63	91194	DKLSSTNQL	TCE (HLA-A*24:02)
sGP; ssGP	130	REL	4.47; 4.43	54.51; 61.10	124–132	91086	AAPDGI**R**GF	TCE (HLA-A*24:02; HLA-B*07:02)
	130, 134	REL	4.47; 4.43	54.51; 61.10	127–135	91188	DGI**R**GFP**R**C	TCE (HLA-A*02:01)
	130, 134, 136	REL	4.47; 4.43	54.51; 61.10	128–136	91371	GI**R**GFP**R**C**R**	TCE (HLA-A*03:01, 11:01; B*07:02)
					129–137	91523	I**R**GFP**R**C**R**Y	TCE (HLA-A*24:02; HLA-B*15:01)
	134, 136	REL	4.47; 4.43	54.51; 61.10	131–139	91360	GFP**R**C**R**YVH	TCE (HLA-A*24:02; HLA-B*07:02, 08:01)
sGP	323, 324, 363	REL	4.47	54.51, 55.76, 55.76	–	–	–	–
Negatively selected amino acid sites								
NP	185	FUBAR	−6.59	0.93	182–190	91828	QGL**I**QYPTA	TCE (HLA-A*02:01, 24:02)
					183–191	20933	GL**I**QYPTAW	TCE (HLA-B*15:01)
					179–187	23528	HAEQGL**I**QY	TCE (HLA-B*35:01)
	341	FEL; FUBAR	−282.57; −5.83	0.03; 0.90	338–346	91260	EQY**Q**QLREA	TCE (HLA-A*02:01)
					335–343	91779	NVGEQY**Q**QL	TCE (HLA-A*02:01, 24:02)
	460	FEL; FUBAR	−213.79; −6.38	0.04; 0.92	451–460	188113	DTTIPDVVV**D**	BCE
	624	FEL; FUBAR	−282.57; −5.92	0.04; 0.91	611–630	188617	YRDHSEKKELPQD**E**QQDQDH	BCE
VP40	9	FUBAR	−5.13	0.90	–	–	**A**	–
VP24	55, 133	FUBAR, FEL	−5.00, −588.12	0.90, 0.03	–	–	**A**, **H**	–

Table 4 continued

Protein	Statistically reliable amino acid sites	Method(s)	dN-dS/ω[a]	*p*-value/Posterior Prob./Bayes factor[b]	Functionally known Epitope length (from-to)	IEDB-Epitope ID[c]	Linear sequence(s)[d]	Function(s)
L	121, 781, 813, 1461, 1625, 1854, 2135	FUBAR	−6.89, −6.92, −6.90, −7.01, −6.86, −6.71, −6.95	0.90, 0.90, 0.91, 0.90, 0.90, 0.90, 0.91	–	–	**V**, **S**, **P**, **K**, **V**, **F**, **F**	–

[a] It was appropriately scaled dN-dS from SLAC/FEL/IFEL/REL of Datamonkey

[b] It shows *p* values of the SLAC/FEL/IFEL (or) the posterior probability of FUBAR method (or) the Bayes factor value of the REL method (the posterior probabilities are included just for reference)

[c] Epitope identification codes of functionally known epitopes that were obtained from IEDB

[d] The corresponding linear amino acid sequences of each functionally known epitope. We have highlighted the amino acids, which were under selection pressures

The presence of both positively and negatively selected amino acid sites in EBOV confirm that the virus evolved in human host under severe immune selection pressures. In natural selection analysis, the significance of one method alone is not sufficient to infer that a given amino acid site underwent either positive or purifying selection pressure; thus, only amino acid sites that have been detected by more than one method are finally considered as highly variable positively or negatively selected sites [8, 24]. In the present study, no positively selected sites have been detected by more than one method. Among the 14 statistically reliable negatively selected sites from NP, VP40, VP24 and L proteins, only 3 sites (341, 460 and 624) of NP protein were identified by more than one method. Hence, these 3 amino acid sites found to be more reliable; thus the corresponding NP gene was used for RNA secondary structure prediction.

RNA secondary structure was predicted for NP gene, which had three synonymous mutations that could be identified by multiple statistical methods. Typically, the protein coding and non-coding regions of viral RNA genome can harbor *cis*-acting replication elements. The nucleotide mutations change Watson–Crick base pairing of the single stranded RNA genome, thus altering the higher order secondary or tertiary RNA structures. It is hypothesized that the synonymous mutations also affect the RNA structural elements without affecting the protein sequence [4]. To test this hypothesis, we predicted the secondary RNA structure of 3 of 11 NP protein-coding genes. The overall comparison results showed that the codons of NPs carrying synonymous mutations altered the RNA secondary structures and optimal energies. Our analysis indicates that the mutated three sites can change the RNA structure that can possibly contribute to increase in viral virulence.

Human immune response against various pathogens depends on T- lymphocytes, specifically CD4+ and CD8+ T cells subsets. T-lymphocytes recognize epitopes of foreign antigens, which are presented by classical human leukocyte antigens (HLAs) [5]. T-cells provide immune protection by identifying TCEs of viruses [6, 29]. Amino acid mutations in many TCEs are expected to be advantageous for viral survival, since, the haplotype of HLA restricts T-cells to recognize TCEs. Apart from natural immune responses, vaccines promote acquired immunity against viral infections. Due to the importance in understanding the function of amino acid sites of EBOV that are under selection pressures, we have aligned these sites to the identical amino acid sites from the functionally known TCEs and BCEs.

The majority of the sites under selection pressure were found to be involved in TCEs than those in BCEs. The mutated sites can provide a selective advantage for virus to

survive in human host. Fourteen amino acid sites engaged in TCEs and BCEs are considered to be appropriate vaccine targets, since these sites have the capability to induce cellular and humoral immune responses. Further experimental validation is required to identify the biological significance of the amino acids under positive or negative selection pressure. Special attention need to be given to the amino acid sites (1 site in sGP and 10 sites in VP40, VP24 and L) that are not mapped as BCE or TCE. Taken together, our results showed that EBOV does not appear to be under diversifying selection pressure, but under purifying selection that appears to be the driving force on their protein coding genes. Results provided an alternative perspective than the current prevailing hypothesis that genes encoding antigens can be highly variable to evade host immunity [5, 9, 16]. Moreover, the synonymous nucleotide changes could also exert modifications to the secondary structure of regulatory cis-RNA sequences. Our analysis provides insight into the importance of host immune selection pressure and replication fitness of EBOV strains involved in the recent outbreak.

Acknowledgments We are grateful to authors and researchers who have made available of all 2014 EBOV genomic sequences in the public database. No funders had a role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Deisy Contreras of Cedars-Sinai Medical Center, Los Angeles for editing this manuscript.

References

1. Arunachalam R. Adaptive evolution of a novel avian-origin influenza A/H7N9 virus. Genomics. 2014;104:545–53.
2. Arunachalam R. Detection of site-specific positive Darwinian selection on pandemic influenza A/H1N1 virus genome: integrative approaches. Genetica. 2013;141:143–55.
3. Becquart P, Mahlakoiv T, Nkoghe D, Leroy EM. Identification of continuous human B-cell epitopes in the VP35, VP40, nucleoprotein and glycoprotein of ebola virus. PLoS ONE. 2014;9:e96360.
4. Chu D, Ren S, Hu S, Wang WG, Subramanian A, Contreras D, Kanagavel V, Chung E, Ko J, Appadorai RSAJ, Sinha S, Jalali Z, Hardy DW, French SW, Arumugaswami V. Systematic analysis of enhancer and critical cis-acting RNA elements in the protein-encoding region of the hepatitis C virus genome. J Virol. 2013;87:5678–96.
5. Comas I, Chakravartti J, Small PM, Galagan J, Niemann S, Kremer K, Ernst JD, Gagneux S. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. Nat Genet. 2010;42:498–503.
6. Cusick MF, Wang S, Eckels DD. In vitro responses to avian influenza H5 by human CD4 T cells. J Immunol. 2009;183:6432–41.
7. de La Vega M-A, Wong G, Kobinger GP, Qiu X. The multiple roles of sGP in Ebola pathogenesis. Viral Immunol. 2014;28:1–7.
8. Delport W, Poon A-FY, Frost SDW, Pond SLK. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics. 2010;2010(26):2455–7.
9. Farci P, Shimoda A, Coiana A, Diaz G, Peddis G, Melpolder JC, Strazzera A, Chien DY, Munoz SJ, Balestrieri A, Purcell RH, Alter HJ. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. Science. 2000;288:339–44.
10. Furuse Y, Shimabukuro K, Odagiri T, Sawayama R, Okada T, Khandaker I, Suzuki A, Oshitani H. Comparison of selection pressures on the HA gene of pandemic (2009) and seasonal human and swine influenza A H1 subtype viruses. Virology. 2010;405:314–21.
11. Geisbert TW, Jahrling PB. Differentiation of filoviruses by electron microscopy. Virus Res. 1995;39:129–50.
12. Harty RN, Brown ME, Wang G, Huibregtse J, Hayes FP. A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. Proc Natl Acad Sci U S A. 2000;97:13871–6.
13. Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. Nat Rev Immunol. 2014;14:36–49.
14. Johnson RF, Bell P, Harty RN. Effect of Ebola virus proteins GP, NP and VP35 on VP40 VLP morphology. Virol J. 2006;3:31.
15. Jun SR, Leuze MR, Nookaew I, Uberbacher EC, Land M, Zhang Q, Wanchai V, Chai J, Nielsen M, Trolle T, Lund O, Buzard GS, Pedersen TD, Wassenaar TM, Ussery DW. Ebolavirus comparative genomics. FEMS Microbiol Rev. 2015;39:764–78.
16. Kawashima Y, Pfafferott K, Frater J, Matthews P, Payne R, Addo M, Gatanaga H, Fujiwara M, Hachiya A, Koizumi H, Kuse N, Oka S, Duda A, Prendergast A, Crawford H, Leslie A, Brumme Z, Brumme C, Allen T, Brander C, Kaslow R, Tang J, Hunter E, Allen S, Mulenga J, Branch S, Roach T, John M, Mallal S, Ogwu A, Shapiro R, Prado JG, Fidler S, Weber J, Pybus OG, Klenerman P, Ndung'u T, Phillips R, Heckerman D, Harrigan PR, Walker BD, Takiguchi M, Goulder P. Adaptation of HIV-1 to human leukocyte antigen class I. Nature. 2009;458:641–5.
17. Kindzelskii AL, Yang Z, Nabel GJ, Todd RF, Petty HR. Ebola virus secretory glycoprotein (sGP) diminishes Fc gamma RIIIB-to-CR3 proximity on neutrophils. J Immunol. 2000;164:953–8.
18. Misasi J, Sullivan NJ. Camouflage and misdirection: the full-on assault of Ebola virus disease. Cell. 2014;159:477–86.
19. Mohamadzadeh M, Chen L, Schmaljohn AL. How Ebola and Marburg viruses battle the immune system. Nat Rev Immunol. 2007;7:556–67.
20. Mohan GS, Li W, Ye L, Compans RW, Yang C. Antigenic subversion: a novel mechanism of host immune evasion by Ebola virus. PLoS Pathog. 2012;8:e1003065.
21. Muhlberger E, Weik M, Volchkov VE, Klenk HD, Becker S. Comparison of the transcription and replication strategies of Marburg virus and Ebola virus by using artiÆcial replication systems. J Virol. 1999;73:2333–42.
22. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Pond SLK, Scheffler K. FUBAR: a fast, unconstrained Bayesian approximation for inferring selection. Mol Biol Evol. 2013;30:1196–205.
23. Platanias LC. Mechanisms of type-I- and type-II-interferon-mediated signalling. Nat Rev Immunol. 2005;5:375–86.
24. Pond SLK, Frost SDW, Grossman Z, Gravenor MB, Richman DD, Brown AJ. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. PLoS Comput Biol. 2006;2:e62.
25. Pond SLK, Frost SDW. Not so different after all: a comparison of methods for detecting amino acid sited under selection. Mol Biol Evol. 2005;22:1208–22.
26. Pond SLK, Murrell B, Fourment M, Frost SDW, Delport W, Scheffler K. A random effects branch-site model for detecting episodic diversifying selection. Mol Biol Evol. 2011;28:3033–43.
27. Reid SP, Leung LW, Hartman AL, Martinez O, Shaw ML, Carbonnelle C, Volchkov VE, Nichol ST, Basler CF. Ebola virus

VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. J Virol. 2006;80:5156–67.
28. Ross AGP, Olveda RM, Yuesheng L. Are we ready for a global pandemic of Ebola virus? Int J Infect Dis. 2014;28:217–8.
29. Suzuki Y. Natural selection on the influenza virus genome. Mol Biol Evol. 2006;23:1902–11.
30. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30:2725–9.
31. Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, Damle R, Sette A, Peters B. The immune epitope database 2. Nucleic Acids Res. 2010;38:D854–62.
32. Yang Z, Delgado R, Xu L, Todd RF, Nabel EG, Sanchez A, Nabel GJ. Distinct cellular interactions of secreted and transmembrane Ebola virus glycoproteins. Science. 1998;279:1034–7.
33. Zuker M. mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003;31:3406–15.